Electronic Supplementary Information

A priori prediction of complex liquid–liquid–liquid equilibria in organic systems using a continuum solvation model

Priyotosh Bairagya [a, #], Debashis Kundu [b, c, #], Tamal Banerjee [c,*]

[a] Department of Chemical Engineering, Jadavpur University, India

[b] Department of Chemical Engineering, Institute of Chemical Technology Marathwada Campus, Jalna

[c] Department of Chemical Engineering, Indian Institute of Technology Guwahati, India

* Corresponding author E-mail: tamalb@iitg.ac.in

# Both the authors have equal contributions

Fax: +91 361 2582291, Tel: +91 361 2582266

Table of Contents

Table S1. Governing equations and universal parameters of the COSMO-SAC (2002) model.

Governing equations of the COSMO-SAC model

sl. no.	description	equation	reference
1.	pair-wise interaction energy ($E_{pair}(\sigma_A, \sigma_B)$)	$E_{pair}(\sigma_A, \sigma_B) = E_{misfit} + E_{hb}$	1
2.	misfit energy (E_{misfit})	$E_{misfit} = (\alpha'/2)(\sigma_A + \sigma_B)^2$	1
3.	h-bonding energy (E_{hb})	$E_{hb} = c_{hb} * \min[0, \max(0, \sigma_{acc} - \sigma_{hb}) * \min(0, \sigma_{don} + \sigma_{hb})]$	1
4.	segment activity coefficient ($\ln(\Gamma_s(\sigma_A))$)	$\ln(\Gamma_s(\sigma_A)) = -\ln\left\{\sum_{\sigma_n} p_s(\sigma_A)\Gamma_s(\sigma_A) * \exp\left[-\frac{E_{pair}(\sigma_A, \sigma_B)}{kT}\right]\right\}$	1
5.	overall activity coefficient ($\ln(\gamma_{i/S})$)	$\ln(\gamma_{i/S}) = n_i \sum_{\sigma_m} p_i(\sigma_A)[\ln(\Gamma_s(\sigma_A)) - \ln(\Gamma_i(\sigma_A))] + \ln(\gamma_{i/S}^{comb})$	1

Universal COSMO-SAC parameters			
sl. no.	description	parameter value	reference
1.	a_{eff} (Å^2)	6.32	2
2.	α' ($kmol * Å^4 mol^{-1} e^{-2}$)	8419	2
3.	c_{hb} ($kcal * Å^4 mol^{-1} e^{-2}$)	75006	2
4.	σ_{hb} ($e * Å^{-2}$)	0.0084	2

Nomenclature:

symbol	**description**
σ_A, σ_B	charge density within the limit $\pm 0.035\ e * Å^{-2}$
α'	misfit energy constant
$\sigma_{acc}, \sigma_{don}$	charge density of acceptor and donor segments respectively
σ_{hb}	hydrogen bonding cut-off value
c_{hb}	hydrogen bonding constant
$p_s(\sigma_A)$	mixture σ-profile
k	Boltzmann constant
$\ln(\Gamma_i(\sigma_A))$	pure-component segment activity coefficient
n_i	cumulative number of surface segments
$\ln(\gamma_{i/S}^{comb})$	Stavermann-Guggenheim combinatorial term

References:

1. S.-T. Lin and S. I. Sandler, Ind. Eng. Chem. Res., 2002, 41, 899-913.

2. A. Bharati, D. Kundu, D. Rabari and T. Banerjee, Phase Equilibria in Ionic Liquid Facilitated Liquid-Liquid Extractions, CRC Press, Boca Raton, 1st edn, 2017.

Table S2. COSMO-SAC predicted liquid phase split percentages and phase fractions calculated by HRL3E algorithm

type of system	system number	temperature (K)	liquid phase split (%)			phase fractions	
			L^T	L^M	L^B	α	β
non-IL systems (ternary)	**1.**	296.15	31.41	29.59	39.00	0.3141	0.4314
			31.50	29.78	38.72	0.3150	0.4347
			31.48	29.70	38.83	0.3148	0.4334
			31.49	29.77	38.74	0.3149	0.4345
			31.56	29.90	38.54	0.3156	0.4369
			31.48	29.74	38.78	0.3148	0.4340
			31.51	29.80	38.69	0.3151	0.4351
	2.	296.15	31.94	32.17	35.89	0.3194	0.4726
	3.	298.15	32.13	32.32	35.56	0.3213	0.4762
		313.15	32.30	31.06	36.64	0.3230	0.4588
		323.15	32.15	30.78	37.07	0.3215	0.4537
		333.15	31.91	30.64	37.45	0.3191	0.4500
		343.15	31.72	30.01	38.27	0.3172	0.4395
		353.15	31.26	29.13	39.61	0.3126	0.4238
		363.15	30.98	28.51	40.50	0.3098	0.4131
	4.	293.15	32.34	31.96	35.71	0.3234	0.4723
	5.	313.15	30.33	39.69	29.98	0.3033	0.5697
		318.15	29.52	40.76	29.72	0.2952	0.5783
		323.15	28.50	42.42	29.09	0.2850	0.5932
IL systems (ternary)	**6.**	298.15	32.30	32.88	34.82	0.3283	0.5245
		348.15	30.32	31.53	38.15	0.3185	0.5834
	7.	298.15	33.22	32.85	33.93	0.3428	0.5204
		348.15	29.21	30.41	40.38	0.3200	0.6058
	8.	298.15	29.83	45.03	25.14	0.2983	0.6417
		348.15	31.06	38.06	30.87	0.3106	0.5521
	9.	298.15	31.25	34.67	34.08	0.3125	0.5044

			31.25	34.80	33.95	0.3125	0.5061
		348.15	50.00	15.31	34.69	0.5000	0.3062
			50.00	15.60	34.40	0.5000	0.3119
non-IL systems (quaternary)	**10.**	293.15	27.60	43.78	28.62	0.2777	0.5965
			26.50	45.21	28.29	0.2791	0.5962
			23.18	50.59	26.24	0.2803	0.5960
			25.19	46.68	28.13	0.2793	0.5995
			23.50	48.52	27.98	0.2808	0.5988
			20.78	51.58	27.63	0.2811	0.6012
	11.	293.15	27.77	43.08	29.15	0.2760	0.6047
			27.91	42.99	29.11	0.2650	0.6151
			28.03	42.89	29.07	0.2318	0.6585
			27.93	43.21	28.87	0.2519	0.6239
			28.08	43.07	28.86	0.2350	0.6343
			28.11	43.22	28.67	0.2078	0.6512

Table S3. Consolidated list of assumptions and boundary conditions applicable to COSMO-SAC, RRL3E and HRL3E models.

List of assumptions

Sl. No.	Model	Assumption	Reference to the section of the main manuscript
1.	COSMO-SAC	Dispersion free energy contribution has been neglected.	Section 2.1
2.	COSMO-SAC	The hypothetical screening medium in COSMO-solvation step has been assumed to be a perfect conductor having infinite dielectric constant	Section 2.1

		($\varepsilon = \infty$). However, for practical solvents the COSMO approximation is accurate to $\pm 0.5\%$ for strong dielectrics for example water and has been chosen as the solvent in our study.[1]	
3.	COSMO-SAC	The σ −profile of the IL has been calculated by linear averaging of cationic and anionic σ-profiles.	Section 3.1
4.	COSMO-SAC	Fractional dissociation of IL components in solution has been neglected since the values of IL dissociation constants (α_0) is empirical in nature and varies with composition and temperature. Complete dissociation of cations and anions is assumed.	Section 3.1
5.	COSMO-SAC	Short and long range weak electrostatic interactions of the IL mixtures using the original Pitzer-Debye-Huckel (PDH) term has been neglected for simplicity and for maintaining thermodynamic consistency.	Section 3.1
6.	RRL3E	The feed composition for each of the LLE subsystems has been chosen as the arithmetic average composition of the top-middle phases for $L^T L^M E$ subsystem and likewise the arithmetic average composition of the middle-bottom phases for the $L^M L^B E$ subsystem.	Section 2.2
7.	HRL3E	The feed composition chosen in this case corresponds to the arithmetic average composition of top, middle and bottom phases.	Section 2.3
8.	HRL3E	Assumption of Sampath and Leipzeiger has been used to simplify the calculation by obtaining α from β and vice-versa using equations 22 and 26.[2]	Section 2.3

List of boundary conditions

Sl. No.	Model	Boundary condition	Reference to the equations

			of the main manuscript
1.	RRL3E	The objective functions of RRL3E and HRL3E algorithms were solved using the modified bisection method which having function bounds corresponding to $f(x_i = 0)$ and $f(x_i = 1)$ where x_i denotes the mole fraction of component i.	Equation 7, 8, 23, 24, 25

References:

1. A. Bharati, D. Kundu, D. Rabari and T. Banerjee, *Phase Equilibria in Ionic Liquid Facilitated Liquid-Liquid Extractions*, CRC Press, Boca Raton, 1st edn, 2017.

2. V. R. Sampath and S. Leipziger, Ind. Eng. Chem. Process Des. Dev., 1985, 24, 652-658.

Table S4. Structures of Ionic Liquid (IL) compounds studied in the present work.

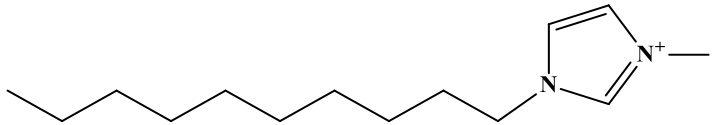
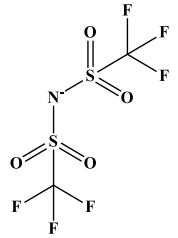
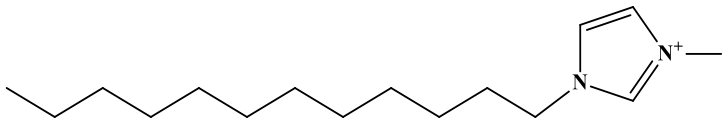
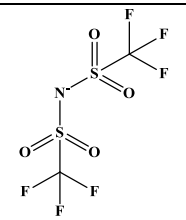
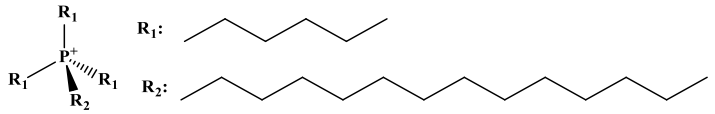
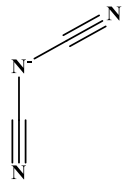
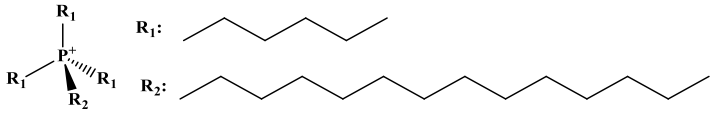
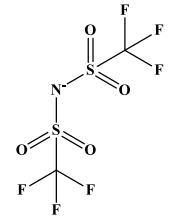
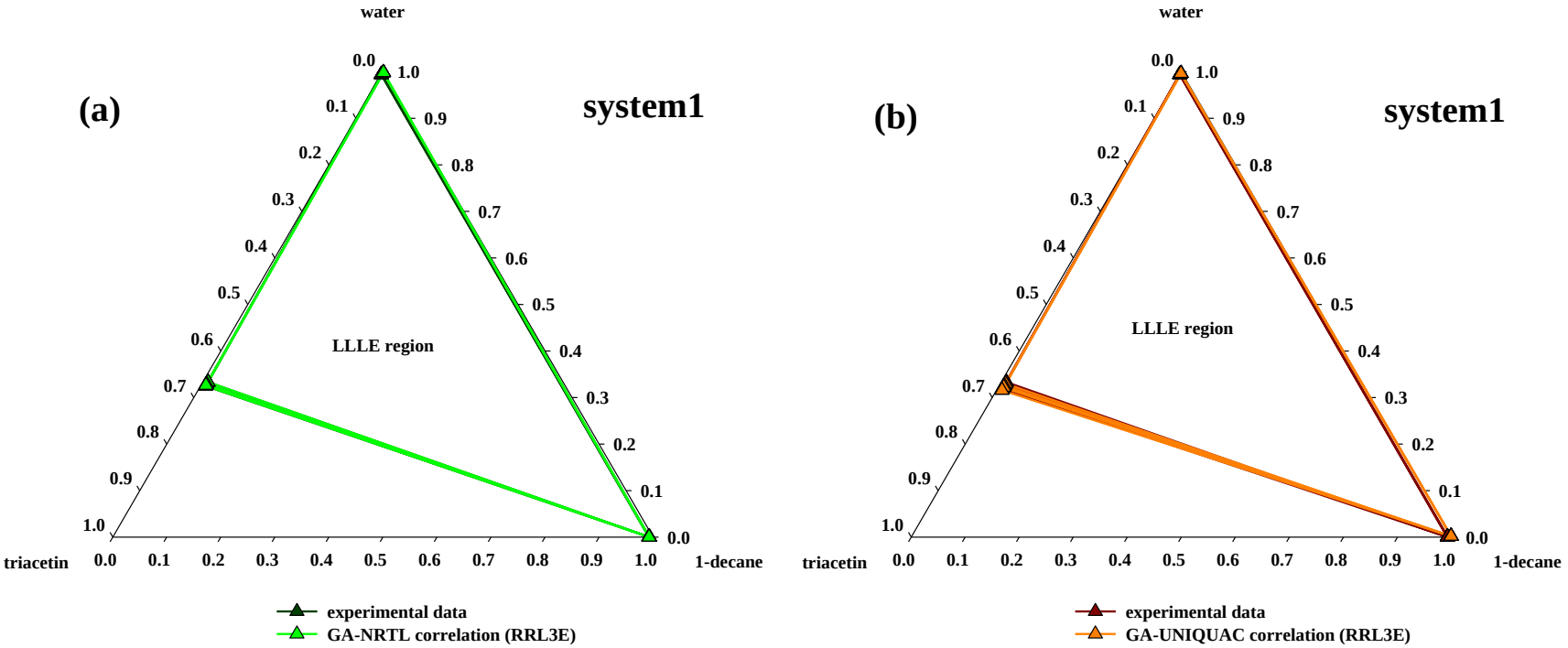
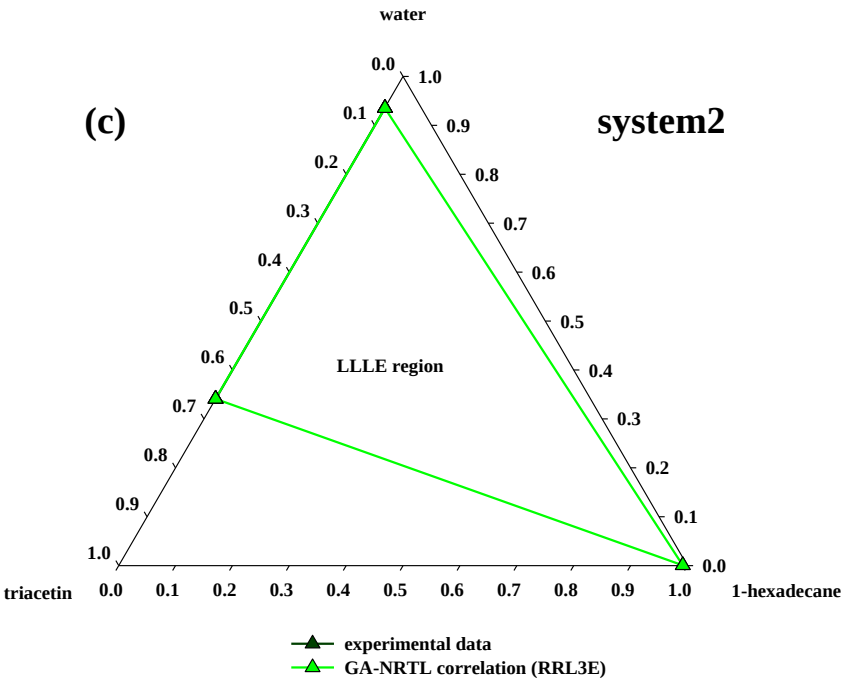
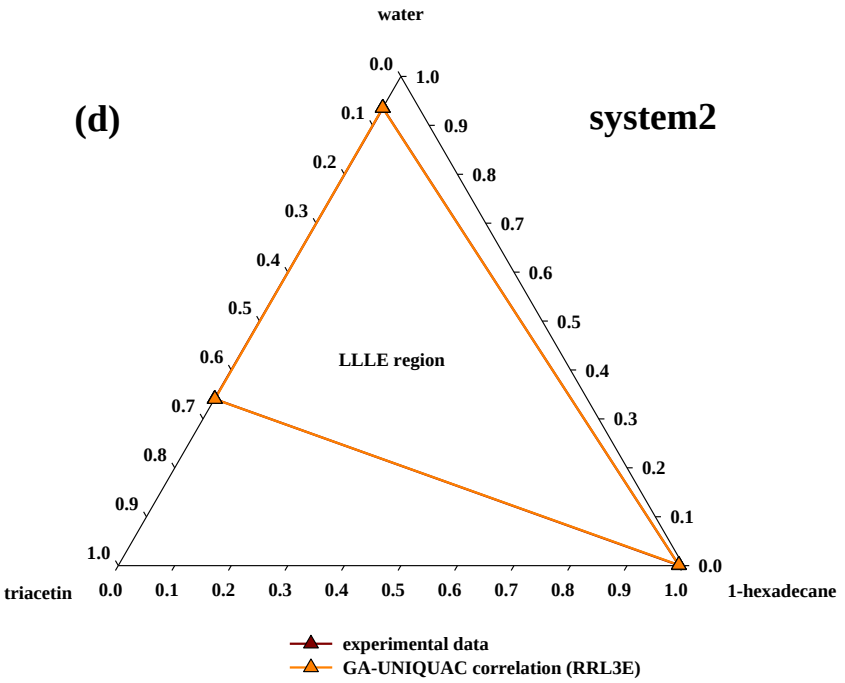
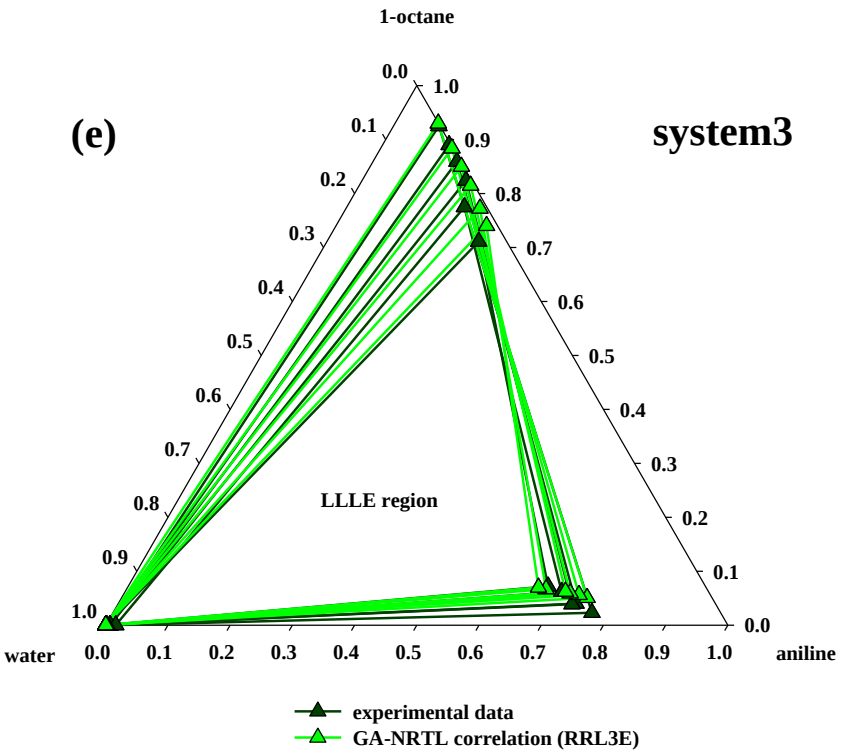
system number	name of the IL	Cation	anion
system6	$[C_{10}mim][NTf_2]$		
system7	$[C_{12}mim][NTf_2]$		
system8	$[P_{66614}][DCA]$		
system9	$[P_{66614}][NTf_2]$		

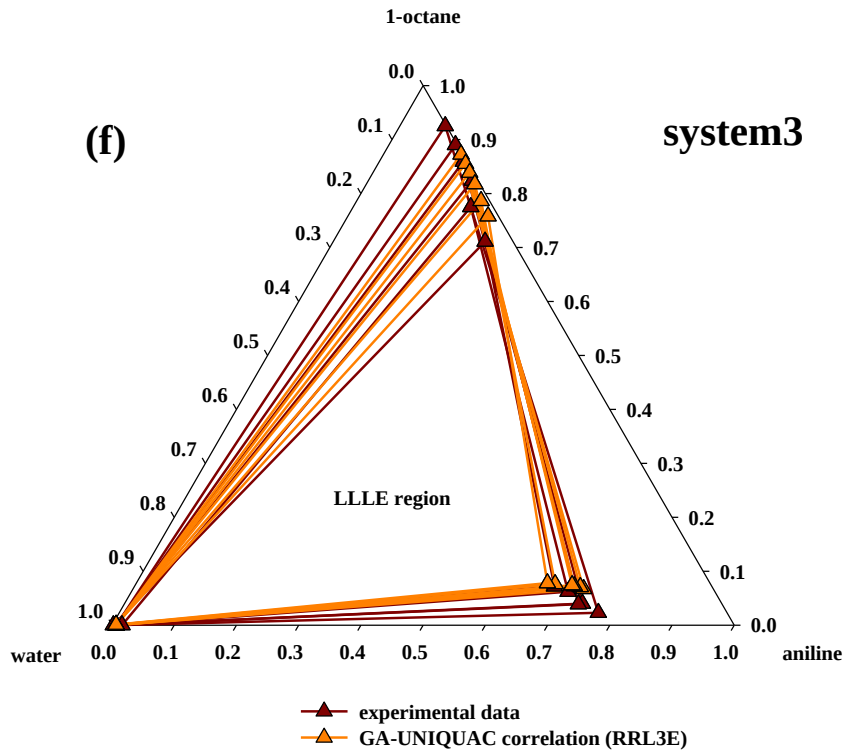
Figure S1. Phase diagrams for ternary LLLE systems correlated using the GA-NRTL (RRL3E) and GA-UNIQUAC (RRL3E) models. (a)-(b) system1, (c)-(d) system2, (e)-(f) system3, (g)-(h) system4, (i)-(j) system5, (k)-(l) system6, (m)-(n) system7, (o)-(p) system8, and (q)-(r) system9.

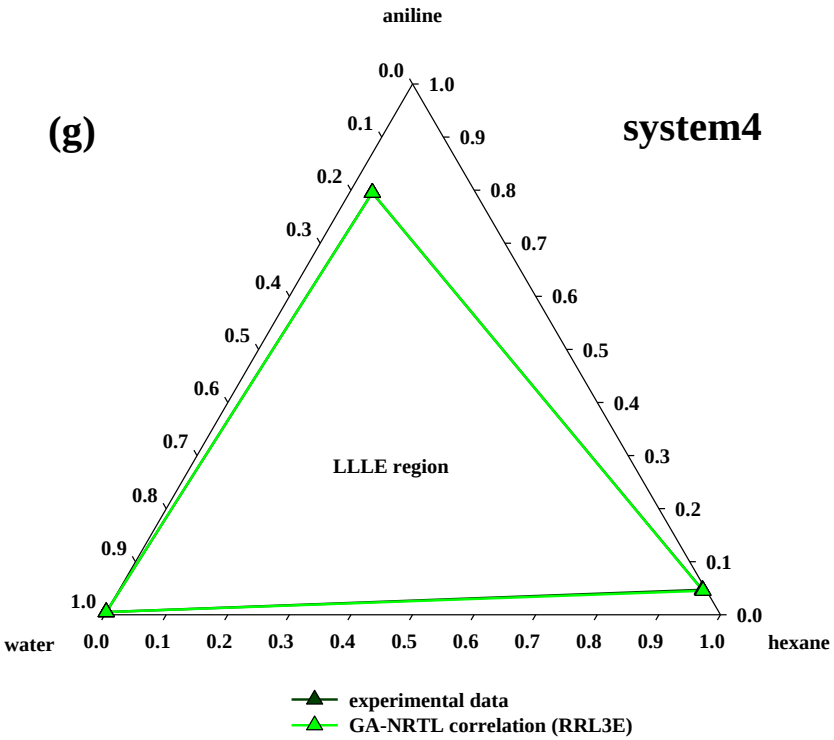


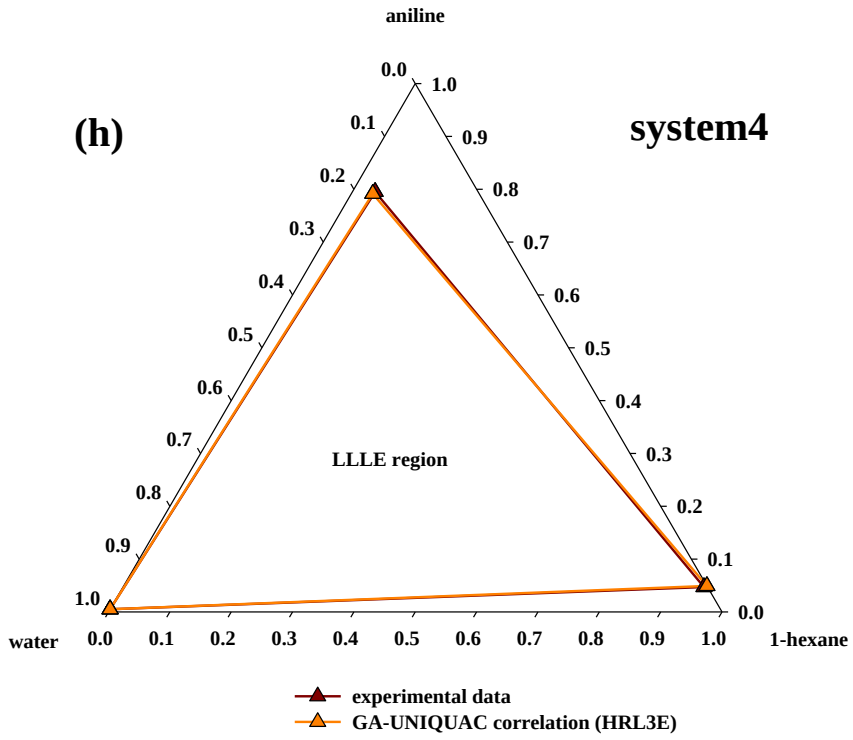

(c)
system2
water
triacetin
1-hexadecane
LLLE region
experimental data
GA-NRTL correlation (RRL3E)

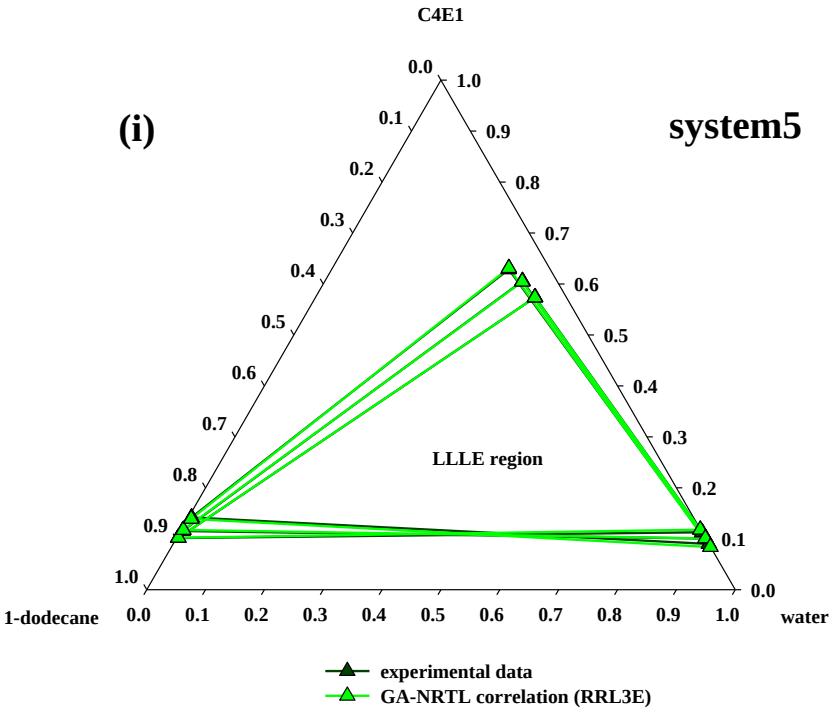

(d)
system2
water
triacetin
1-hexadecane
LLLE region
experimental data
GA-UNIQUAC correlation (RRL3E)

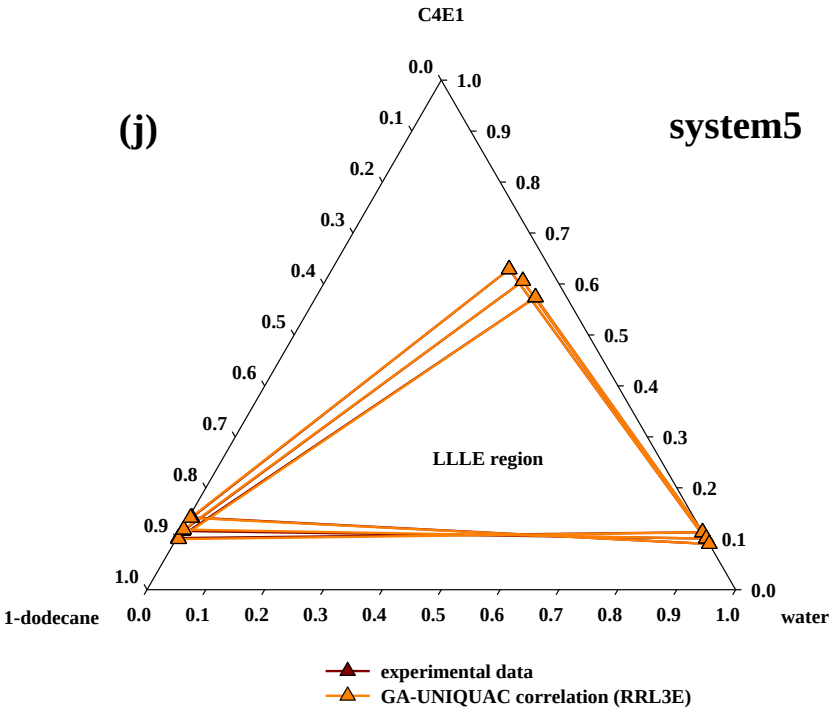

(e)
system3
1-octane
water
aniline
LLLE region
experimental data
GA-NRTL correlation (RRL3E)

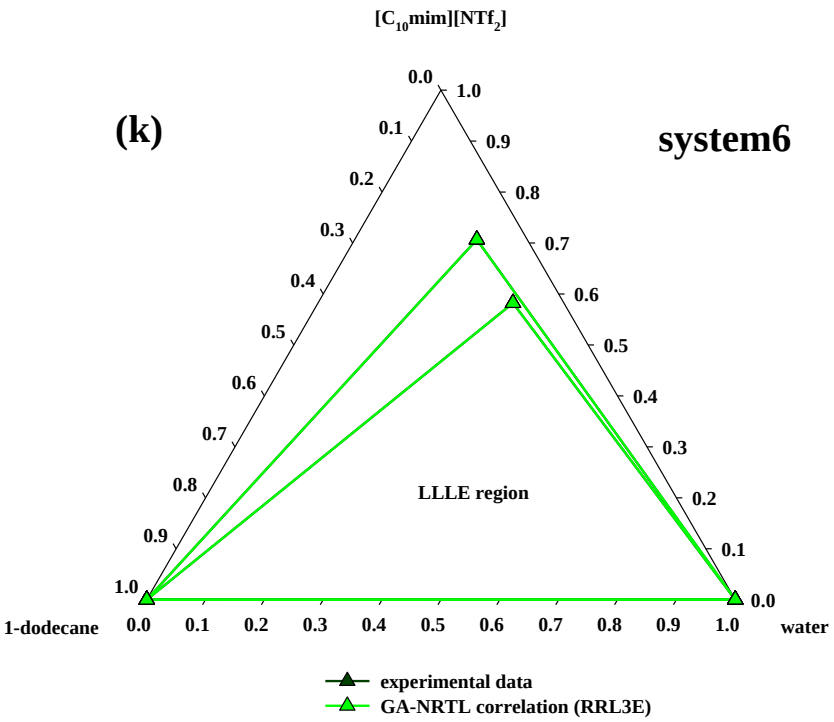

(f)
system3
1-octane
water
aniline
LLLE region
experimental data
GA-UNIQUAC correlation (RRL3E)

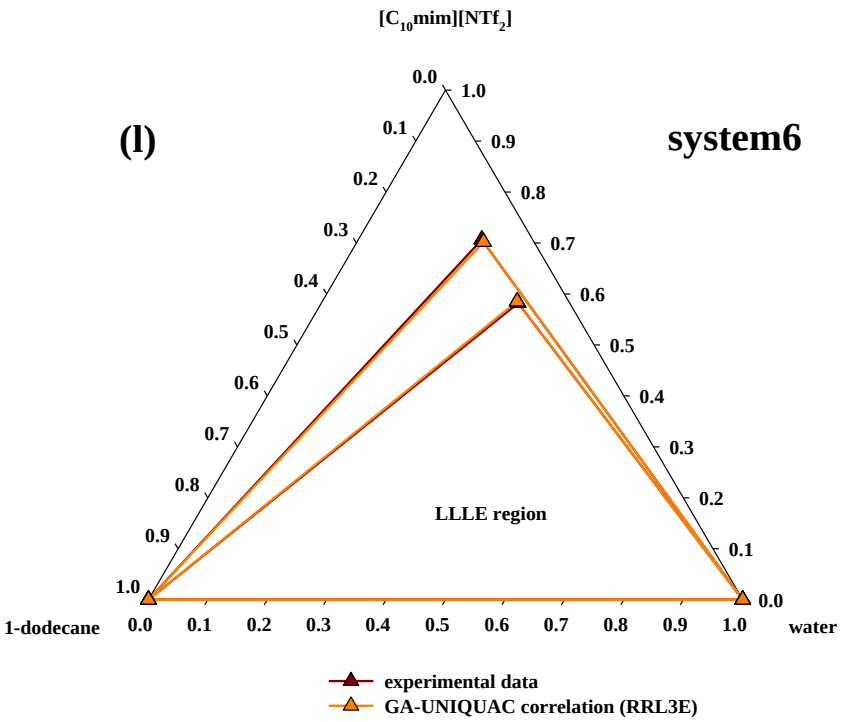

(g)
system4
aniline
water
hexane
LLLE region
experimental data
GA-NRTL correlation (RRL3E)

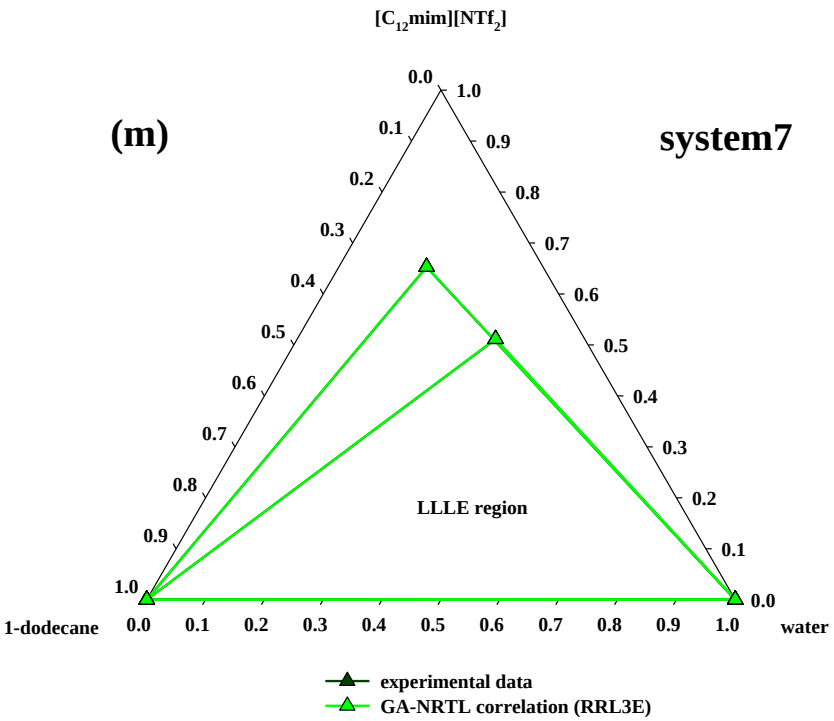

(h)
system4
aniline
water
1-hexane
LLLE region
experimental data
GA-UNIQUAC correlation (HRL3E)

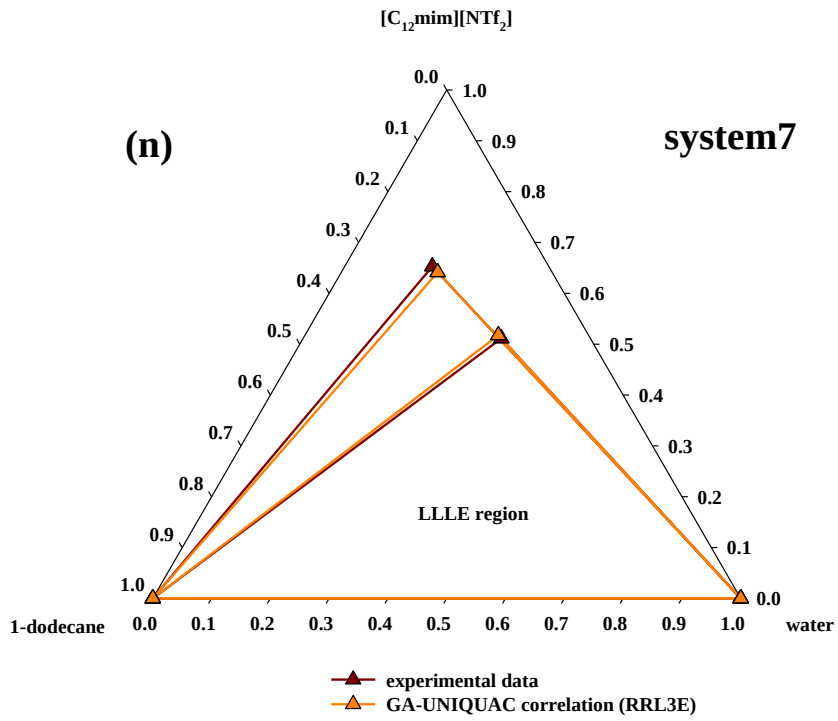

(i)
system5
C4E1
1-dodecane
water
LLLE region
experimental data
GA-NRTL correlation (RRL3E)

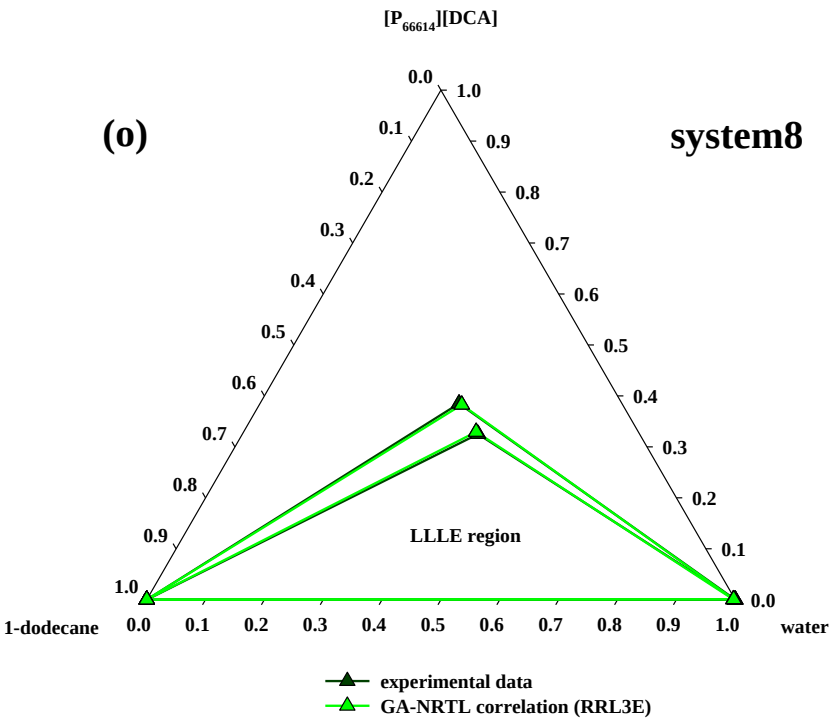

(j)
system5
C4E1
1-dodecane
water
LLLE region
experimental data
GA-UNIQUAC correlation (RRL3E)

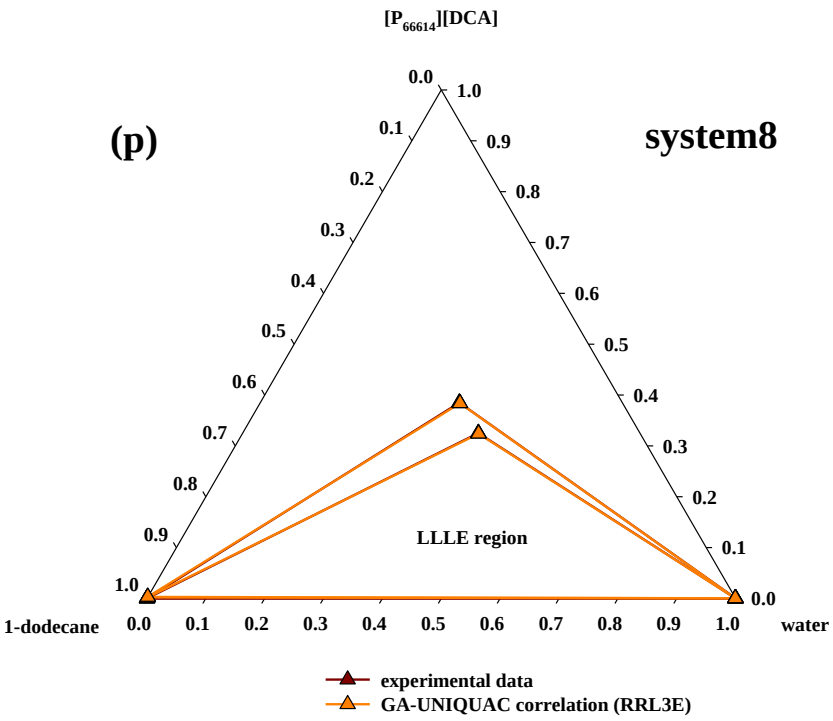

(k)
system6
[C10mim][NTf2]
LLLE region
1-dodecane
water
experimental data
GA-NRTL correlation (RRL3E)

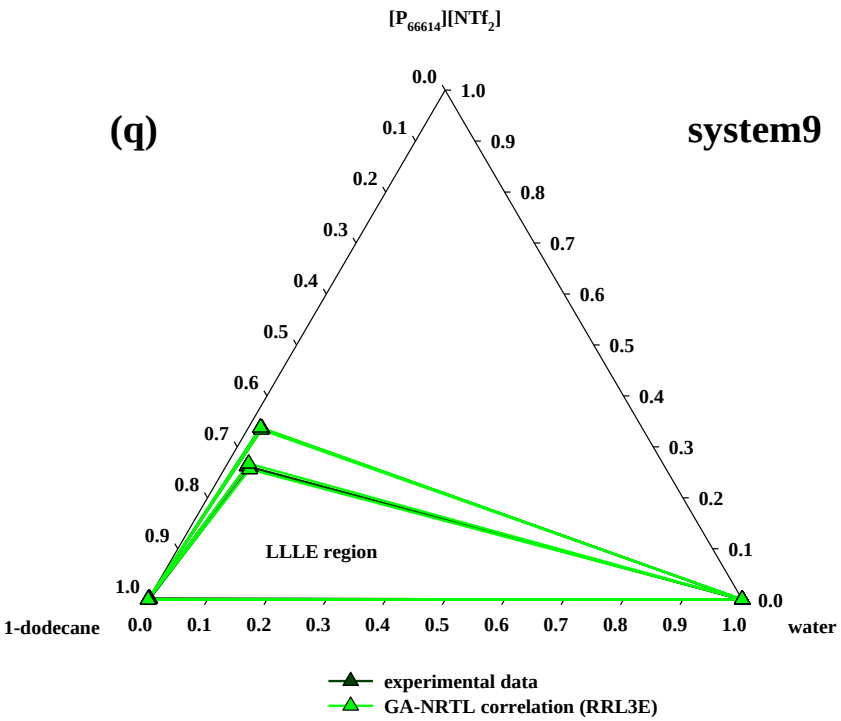

(l)
system6
[C10mim][NTf2]
LLLE region
1-dodecane
water
experimental data
GA-UNIQUAC correlation (RRL3E)


(m)
system7
[C12mim][NTf2]
LLLE region
1-dodecane
water
experimental data
GA-NRTL correlation (RRL3E)


(n)
system7
[C12mim][NTf2]
LLLE region
1-dodecane
water
experimental data
GA-UNIQUAC correlation (RRL3E)


(o)
system8
[P66614][DCA]
LLLE region
1-dodecane
water
experimental data
GA-NRTL correlation (RRL3E)


(p)
system8
[P66614][DCA]
LLLE region
1-dodecane
water
experimental data
GA-UNIQUAC correlation (RRL3E)


(q)
system9
[P66614][NTf2]
LLLE region
1-dodecane
water
experimental data
GA-NRTL correlation (RRL3E)

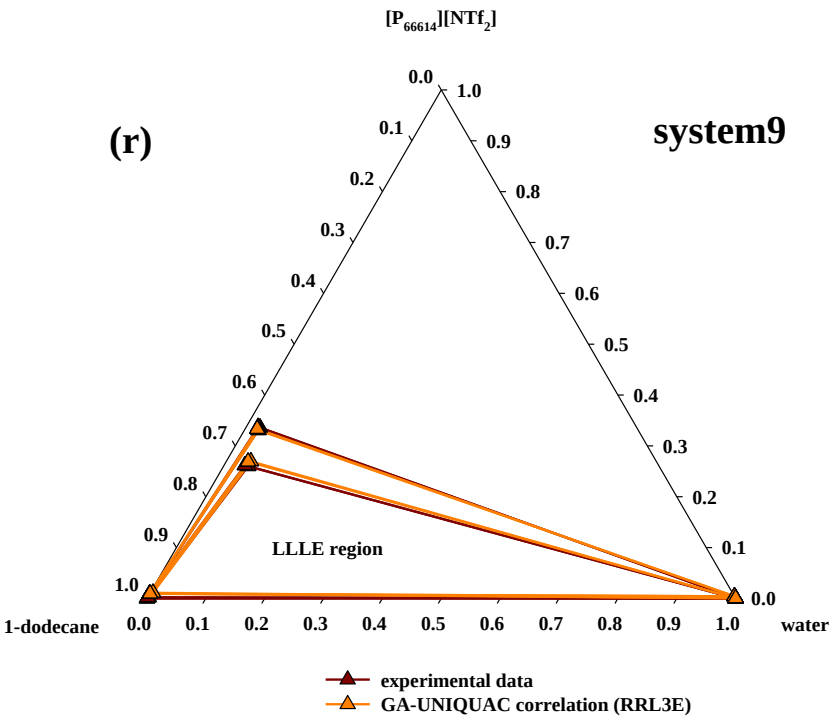
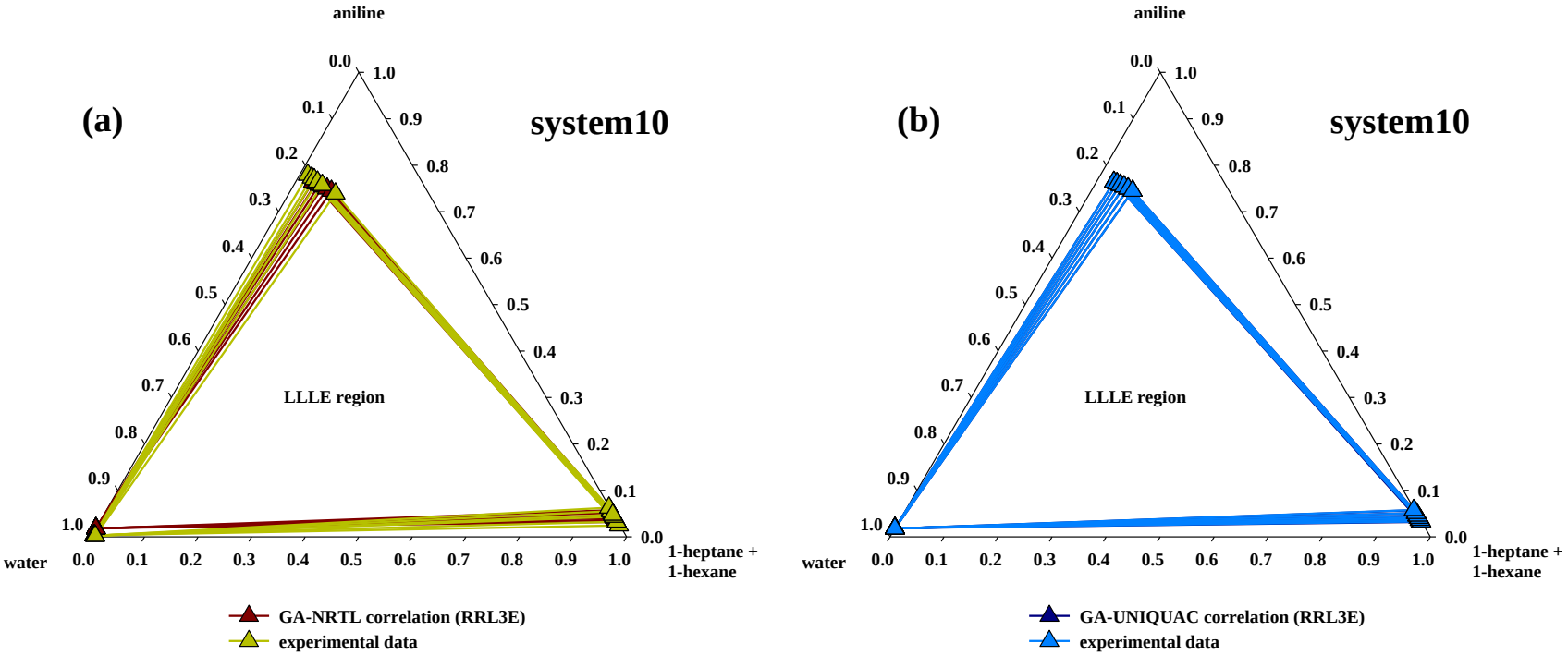
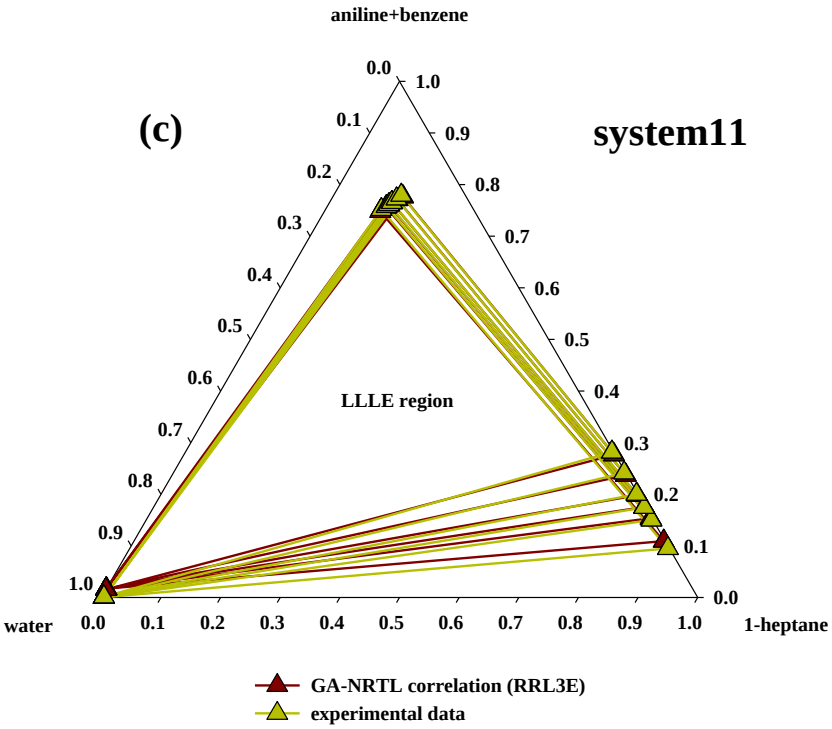

(r)
system9
[P66614][NTf2]
LLLE region
1-dodecane
water
experimental data
GA-UNIQUAC correlation (RRL3E)

Figure S2. Phase diagrams for quaternary LLLE systems correlated using the GA-NRTL (RRL3E) and GA-UNIQUAC (RRL3E) models. (a)-(b) system10; (c)-(d) system11.




(c)
system11
aniline+benzene
water
1-heptane
LLLE region
GA-NRTL correlation (RRL3E)
experimental data

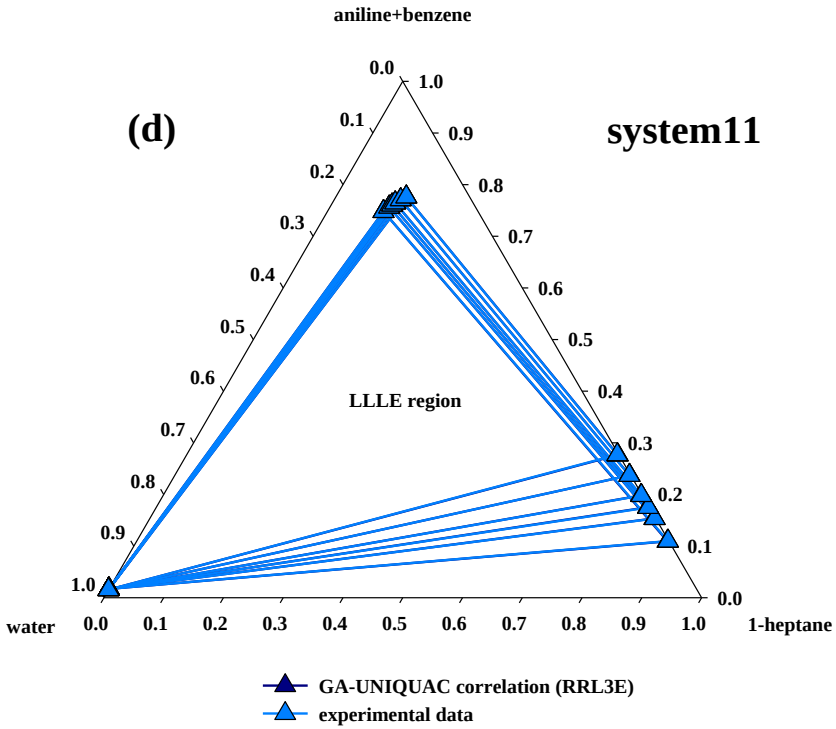

(d)
system11
aniline+benzene
water
1-heptane
LLLE region
GA-UNIQUAC correlation (RRL3E)
experimental data

Table S5. Binary interaction parameters for GA-NRTL correlations of ternary LLLE systems using the RRL3E algorithm.							
type	**system**	**top-middle phase interactions ($L^T L^M E$)**			**middle-bottom phase interactions ($L^M L^B E$)**		
		$i-j$	τ_{ij}	τ_{ji}	$i-j$	τ_{ij}	τ_{ji}
ternary non-IL systems	**system1**	1 – 2	2.877	17.531	1 – 2	2.853	0.739
		1 – 3	1.627	11.446	1 – 3	17.278	9.217
		2 – 3	15.940	18.184	2 – 3	-2.321	15.349
	system2	1 – 2	20.000	4.384	1 – 2	10.231	0.592
		1 – 3	11.359	-10.000	1 – 3	19.883	10.186
		2 – 3	4.384	3.573	2 – 3	19.589	5.167
	system3	1 – 2	3.277	15.031	1 – 2	4.153	9.963
		1 – 3	2.441	2.715	1 – 3	-0.611	15.595
		2 – 3	-1.608	-8.523	2 – 3	11.157	17.907
	system4	1 – 2	5.793	7.704	1 – 2	-5.959	11.753
		1 – 3	0.819	-0.172	1 – 3	1.080	19.943
		2 – 3	-3.242	8.017	2 – 3	18.995	3.558
	system5	1 – 2	20.000	-1.360	1 – 2	18.070	3.083
		1 – 3	15.645	9.381	1 – 3	-8.669	-0.533
		2 – 3	2.277	0.400	2 – 3	-17.875	20.000
ternary IL systems	**system6**	1 – 2	1.866	-5.710	1 – 2	14.942	10.923
		1 – 3	11.025	18.050	1 – 3	15.121	-1.609
		2 – 3	-1.525	1.420	2 – 3	-4.675	10.487

	system7	1 – 2	7.584	2.416	1 – 2	17.050	19.985
		1 – 3	10.448	12.187	1 – 3	13.716	3.809
		2 – 3	18.111	9.790	2 – 3	9.319	9.573
	system8	1 – 2	-9.297	-2.071	1 – 2	-0.033	1.270
		1 – 3	17.427	6.290	1 – 3	6.869	-2.688
		2 – 3	19.486	6.747	2 – 3	4.922	-9.885
	system9	1 – 2	-7.527	12.649	1 – 2	5.613	7.069
		1 – 3	5.697	7.157	1 – 3	11.382	0.948
		2 – 3	5.476	19.655	2 – 3	19.993	11.416

Table S6. Binary interaction parameters for GA-UNIQUAC correlations of ternary LLLE systems using the RRL3E algorithm.

type	system	top-middle phase interactions ($L^T L^M E$)			middle-bottom phase interactions ($L^M L^B E$)		
		$i-j$	A_{ij}/K	A_{ji}/K	$i-j$	A_{ij}/K	A_{ji}/K
ternary non-IL systems	system1	1 – 2	2556.23	822.44	1 – 2	3261.45	43.41
		1 – 3	328.18	-87.95	1 – 3	1346.42	-588.40
		2 – 3	-285.29	1032.38	2 – 3	242.22	1525.44
	system2	1 – 2	9158.78	970.43	1 – 2	-290.57	1590.96
		1 – 3	1621.36	-425.70	1 – 3	-1340.45	4163.49
		2 – 3	1032.25	7173.55	2 – 3	790.73	4413.15
	system3	1 – 2	1239.51	-88.54	1 – 2	-181.76	-7730.24
		1 – 3	2126.64	17.05	1 – 3	5389.07	-5703.37
		2 – 3	1543.12	761.58	2 – 3	2781.47	-762.50
	system4	1 – 2	28.96	1462.04	1 – 2	-647.02	-3847.17
		1 – 3	9885.08	6949.34	1 – 3	2430.00	-4306.51
		2 – 3	4506.08	137.26	2 – 3	4120.71	584.35
	system5	1 – 2	7661.72	2332.62	1 – 2	-8.06	1741.45
		1 – 3	7804.89	3222.28	1 – 3	-4687.49	6135.59
		2 – 3	-205.75	540.74	2 – 3	-9615.66	-542.09
ternary IL systems	system6	1 – 2	-9994.99	3579.06	1 – 2	-589.16	9936.29
		1 – 3	461.04	6657.22	1 – 3	-750.44	4240.27
		2 – 3	-647.71	-8025.58	2 – 3	-4254.16	-9788.90

	system7	1 – 2	7561.09	9852.12	1 – 2	3804.88	8932.18
		1 – 3	474.41	6978.03	1 – 3	-493.94	8827.26
		2 – 3	-310.20	3902.40	2 – 3	-3045.70	1148.20
	system8	1 – 2	-1596.01	-8590.16	1 – 2	-769.01	7640.93
		1 – 3	5371.79	395.08	1 – 3	-7349.61	6952.62
		2 – 3	2212.12	4229.56	2 – 3	-7536.29	-1644.00
	system9	1 – 2	714.44	712.15	1 – 2	218.10	6544.45
		1 – 3	9083.96	8478.15	1 – 3	461.26	6606.71
		2 – 3	690.98	87.46	2 – 3	-29.81	-210.71

Table S7. Binary interaction parameters for GA-NRTL correlations of quaternary LLLE systems using the RRL3E algorithm.										
system number	**top-middle phase interactions ($L^T L^M E$)**					**middle-bottom phase interactions ($L^M L^B E$)**				
10.	A_{ij}/K	1	2	3	4	A_{ij}/K	1	2	3	4
	1	**0.0**	14.35	-9.66	20.00	1	**0.0**	17.69	14.28	20.00
	2	16.93	**0.0**	-15.50	2.36	2	-13.98	**0.0**	17.37	7.86
	3	11.23	9.85	**0.0**	16.83	3	16.61	9.13	**0.0**	4.22
	4	5.84	0.27	19.99	**0.0**	4	-7.87	-11.31	0.88	**0.0**
	f_{min}	-6.49E-04				f_{min}	-2.85E-03			
11.	A_{ij}/K	1	2	3	4	A_{ij}/K	1	2	3	4
	1	**0.0**	19.97	3.69	-11.99	1	**0.0**	5.38	3.10	19.59
	2	19.41	**0.0**	1.83	-9.83	2	-18.44	**0.0**	13.85	9.77
	3	-19.77	-13.70	**0.0**	-11.21	3	-19.97	0.61	**0.0**	7.61
	4	-13.31	-18.96	4.44	**0.0**	4	9.38	20.00	11.99	**0.0**
	f_{min}	-4.79E-03				f_{min}	-2.47E-03			

Table S8. Binary interaction parameters for GA-UNIQUAC correlations of quaternary LLLE systems using the RRL3E algorithm.

system number	top-middle phase interactions ($L^T L^M E$)					middle-bottom phase interactions ($L^M L^B E$)				
10.	A_{ij}/K	*i*=1	2	3	4	A_{ij}/K	1	2	3	4
	j=1	**0.0**	-964.05	-1520.39	-1560.62	1	**0.0**	97.24	5442.42	-2144.77
	2	5842.98	**0.0**	-4088.06	-1424.07	2	-8425.09	**0.0**	1580.12	-3505.88
	3	7332.91	-2041.80	**0.0**	-1684.76	3	-4370.01	290.00	**0.0**	9979.13
	4	9998.62	3328.13	1021.18	**0.0**	4	-7402.54	-241.32	14533.80	**0.0**
	f_{min}	-1.34E-03				f_{min}	-1.01E-05			
11.	A_{ij}/K	*i*=1	2	3	4	A_{ij}/K	1	2	3	4
	j=1	**0.0**	-1318.86	-7428.00	-7519.78	1	**0.0**	9987.00	-1755.15	227.69
	2	6808.32	**0.0**	-6080.78	-4390.40	2	-658.44	**0.0**	-7814.50	2411.76
	3	5513.04	-1266.92	**0.0**	9999.89	3	-1538.67	3047.43	**0.0**	-226.91
	4	2621.80	-2375.99	7200.43	**0.0**	4	434.55	13264.05	-236.54	**0.0**
	f_{min}	-4.63E-04				f_{min}	-1.56E-05			

Table S9. Observed *rmsd* (%) for GA-NRTL (RRL3E) model correlations					
type of system	**system number**	**Individual phase-wise *rmsd* values (%)**			**overall *rmsd* (%)**
		top	**middle**	**bottom**	
non-IL systems (ternary)	**1.**	0.040	0.036	0.428	0.249
	2.	0.003	0.002	0.022	0.013
	3.	3.180	1.705	0.654	2.117
	4.	0.149	0.025	0.001	0.087
	5.	0.168	0.106	0.374	0.245
IL systems (ternary)	**6.**	0.000	0.000	0.003	0.002
	7.	0.000	0.000	0.002	0.001
	8.	0.012	0.405	0.251	0.275
	9.	0.230	0.232	0.056	0.135
non-IL systems (quaternary)	**10.**	0.914	1.350	1.108	0.986
	11.	0.482	0.358	1.183	0.663
average			**ternary systems**		**0.347**
			quaternary systems		**0.825**
			overall		**0.434**

Table S10. Observed *rmsd* (%) for GA-UNIQUAC (RRL3E) model correlations					
type of system	**system number**	**Individual phase-wise *rmsd* values (%)**			**overall *rmsd* (%)**
		top	**middle**	**bottom**	
non-IL systems (ternary)	**1.**	0.024	0.002	0.003	0.014
	2.	0.015	0.002	0.002	0.009
	3.	4.928	1.358	0.380	2.960
	4.	0.000	0.000	0.000	0.000
	5.	0.122	0.000	0.000	0.071
IL systems (ternary)	**6.**	0.001	0.001	0.003	0.002
	7.	0.001	0.001	0.002	0.001
	8.	0.000	0.001	0.000	0.001
	9.	0.045	0.002	0.001	0.026
non-IL systems (quaternary)	**10.**	0.087	0.000	0.000	0.044
	11.	0.094	0.000	0.000	0.047
average			**ternary systems**		**0.343**
			quaternary systems		**0.045**
			overall		**0.288**

Table S11. UNIQUAC structural parameters from PCM calculations					
type	**name**	**GEPOL volume ($Å^3$)**	**GEPOL surface area ($Å^2$)**	$r_{predicted}$	$q_{predicted}$
IL components	[C_{10}mim]	250.827	331.109	9.9573	7.9759
	[C_{12}mim]	284.367	374.409	11.2887	9.0190
	[P_{6614}]	573.534	757.169	22.7680	18.2391
	[DCA]	55.390	81.086	2.1989	1.9532
	[NTf_2]	151.224	202.428	6.0032	4.8762
Non-IL components	aniline	94.885	125.281	3.7667	3.0178
	1-hexane	111.507	155.215	4.4266	3.7389
	1-heptane	128.140	176.567	5.0869	4.2532
	1-dodecane	213.095	287.406	8.4594	6.9232
	1-octane	144.764	197.814	5.7468	4.7651
	phenol	91.706	120.171	3.6405	2.8947
	triacetin	195.538	258.987	7.7624	6.2386
	1-decane	178.007	240.383	7.0665	5.7905
	1-hexadecane	277.772	368.181	11.0269	8.8690
	benzene	84.122	110.837	3.3394	2.6699
	C4E1	127.917	177.469	5.0780	4.2750
	water	19.105	35.964	0.7584	0.8663